

GIS NEWS AND INFORMATION

June 1995 (No. 4)

Public Health GIS Conference and Training Opportunities

CDC/ATSDR

1. The National Center for Health Statistics will host the 1995 Public Health Conference on Records and Statistics (PHCRS), July 17-20, at the Mayflower Hotel in Washington, D.C. A session entitled "Public Health Applications Using Geographic Information Systems" (see April, GIS Communications) will be held Monday, **July 17**, from 3:30-5:00 p.m. The conference is free. For conference information, please contact Barbara Hetzler (301) 436-7122, ext. 148, in Hyattsville, MD.

2. Given the strong staff turnout for the GIS Short Course at the 1995 CDC and ATSDR Symposium on Statistical Methods, taught by Dr. Charles Croner, CDC NCHS, a sequel GIS program focusing on advanced spatial statistical techniques is planned for **September 12-13**. Dr. Luc Anselin, Professor of Economics, Regional Research Institute, West Virginia University and formerly with the National Center for Geographic Information and Analysis (NCGIA), is a nationally recognized geographer and economist for his innovative work in spatial data analysis and modeling. His course will cover the basics of spatial stochastic processes (spatial scale, spatial weights, spatial lags, spatial autocorrelation, spatial filtering, spatial dependence in regression residuals, etc.) and their application using GIS. He will use Arc/View and his own SpaceStat software to demonstrate these principles. Please reserve these dates. Information on attending will be forthcoming from **Jay Smith** and **Don Betts** in the near future.

3. NCHS will offer a 1-day abbreviated workshop by Dr. Anselin similar in content to that offered in Atlanta. It is entitled "Exploratory Spatial Data Analysis Using Geographic Information Systems (GIS)" and is tentatively scheduled for **September 21**, in Hyattsville, MD. Please contact **Chuck Croner** for information.

Non CDC/ATSDR

4. Announcing two summer GIS workshops for instructors and practitioners in public health: "Improving Public Health Using Geographic Information and Analysis", **August 10 - 12** and **August 17 - 19**, Iowa City. The instructors are faculty members at The University of Iowa-Gerard Rushton, Ph.D., Department of Geography, Marc P. Armstrong, Ph.D. Department of Geography, Charles F. Lynch, MD, Ph.D. Department of Preventive Medicine, and James Rohrer, Ph.D. Graduate Program in Hospital and Health Administration. These workshops are supported by a grant from The U.S. Department of Education (FIPSE). Each 3-day workshop is intended for public health teachers and practitioners who have responsibilities for disease surveillance and who wish to learn more about how geographic information systems (GIS) are used in public health. The purpose of these workshops is to provide participants with the ability to measure the geographic rates of disease incidences and other health-related indicators for small areas, to assess the relationships between geographic patterns of disease and socio-economic conditions of populations, to evaluate the geographic accessibility of defined populations to health resources, and to determine appropriate locations for health resources.

Workshop sessions will be held in The University of Iowa instructional computing facility in the newly completed Pappajohn Building on the main campus of The University of Iowa in Iowa City, Iowa. Each student will have access to a networked personal computer.

Participants will learn how to use GIS to perform a detailed analysis of health data and will learn how to: a. acquire digital road maps of local areas from public domain sources or enhanced products from private vendors, b. acquire software to match addresses stored in health files to the digital map, c. compute and map the geographic patterns of disease incidence and to determine the spatial relationship between rates of disease incidence and socio-economic data from sources such as the U.S. Census, d. make tests of statistical significance for geographical patterns of diseases, e. evaluate the geographical pattern of health facilities in relation to need, and f. evaluate alternative locations for appropriate health services development.

Participants should be doctoral students, instructors in schools or departments of public health or epidemiology but are not expected to have any prior knowledge of geographic information systems. The registration fee of \$475 will be waived for qualified applicants from U.S. educational and public health institutions. Registration materials and more information can be obtained from: The University of Iowa, Center for Conferences and Institutes, The University of Iowa, 249 Memorial Union, Iowa City, Iowa 52242-1317, (319) 335-3231 or by E-Mail from gerard-rushton@uiowa.edu

5. The Department of Geography and Environmental Planning, Towson State University, Towson, Maryland, recently

convened (June 7-8) its eighth annual GIS conference. A session on Epidemiology/Risk Assessment included papers "Customizing GIS applications for hazardous waste environmental epidemiologic analyses" (David Padgett, Austin Peay State U.), "Spatiotemporal patterns of Lyme disease risk in Maryland using GIS" (Greg Glass, Johns Hopkins U.), and "Radon hazard-where and why? Vulnerability analysis using GIS" (Kent Barnes, Towson State U. and Charles Geiger, Millersville, U.).

News from CDC/ATSDR GIS USERS

(Please communicate directly with your colleagues on any issues)

General News

1. From **John Odencrantz**: Regarding **Scott Hendricks'** query (April, GIS Communication), auto-logistic models were the topic of my PhD dissertation. I have already informed him of that, but if there are others who have similar interests or questions, this is to let you know. **Editor**: Many thanks to John for his willingness to assist GIS Users in auto-logistic techniques.

2. From **Bob Bernstein**: With technical assistance from CDC/EPO, and through a series of consultations for USAID, the World Bank, and UNDP in Indonesia (the 4th most populous country in the world), I have been successful in demonstrating how the public domain CDC EpiMap software can be used at national, province, and district level for planning, monitoring, and evaluating health sector projects and programs, and for policy dialogue to improve data collection and use in public health surveillance and decision making. Based on these demonstrations, the Ministry of Health has officially adopted

EpiMap for use at province and district-level in Indonesia. While EpiMap is not a "true" GIS (e.g., it is not possible to overlay data of different types on the same boundary), the ability to quickly and easily place several maps side-by-side as well as the ease of linking EpiMap with EpiInfo databases that are widely used for surveillance and surveys in Indonesia convinced users of expensive commercial GIS software that EpiMap was the most suitable software for use below national levels. Despite the complicated nature of the terrain and political geography of Indonesia (an archipelago of more than 13,000 islands of which more than 6,000 are populated), I was able to use a mouse to prepare more than 300 district-level boundary files for Indonesia's 27 provinces by tracing the boundaries from transparencies of photocopied maps purchased from the Indonesia Census Bureau.

3. From **Charles Irland** (505) 837-4226: My inquiry is about GIS software and available electronic maps. Dr. Ralph Bryan (CDC), Dr. Cheek (IHS) and I (all in Albuquerque, NM) are going to use GIS software for some environmental analysis on hantavirus cases. We are in the initial stages of determining what S/W will be required. The packages that I am aware of are EpiMap, AtlasGIS, MapInfo, ArcInfo (and related ESRI products), IDRISI and ERDAS Imagine. I do not know of all the capabilities (and limitations) of the different packages. It is anticipated that we want to use ecological, climatic, topographical and census maps (raster and vector based) to do some correlation studies, spatial analysis and maybe some prediction modeling. Any direction that we take should be along the main stream of where CDC is headed. What are other groups using? I believe the most important chunk of information that I am looking for has to do with electronic maps. Is there anyone or place

that is cataloging the maps that are available internal and external to CDC? This would also include point of contacts (POCs) for maps external to CDC. I realize that there are diverse external sources. I have started to make inquiries to USGS, EROS, USFS, National Climatic Data Center and the Army Corps of Engineers. My next step is to start cruising through the Internet. I have an ESRI Digital Chart of the World available at UNM. Does anyone have a POC at the Defense Mapping Agency and are there any special procedures?

4. From **Alan Hightower** (to be presented later this year with several coauthors): Differential Global Positioning System (DGPS) technology was used to map four adjacent villages participating in the Malaria Cohort project in Western Kenya for use in a Geographic Information System (GIS). This technology, which requires simultaneous use of two GPS units reduces errors of measurement to less than five meters versus 100 meters for readings from a single GPS unit. The longitudinal nature of the Malaria Cohort Study and the fact that all data are collected with household identifiers presents unique opportunities for multi-disciplinary spatial analyses. The longitude, latitude, and altitude of study households, mosquito breeding sites, local health care clinics, medicine stores, roads, rivers, the shoreline of Lake Victoria, and other features was computed and entered into a database file. Study databases from a variety of disciplines were then linked to the map feature databases via GIS software. Buffer zones were computed around the lake shore and mosquito breeding sites to summarize and compare data for households inside and outside the buffers.

We will review how DGPS technology works, how it was used here, steps taken to verify the accuracy of the points, personnel and

time commitments needed, and how the DGPS readings are used to create a GIS map file. Steps needed to link GIS maps to study databases will be presented. Preliminary data from two studies (entomology and childhood mortality) will be used to show how GIS is being used to investigate spatial hypotheses.

Technical News

5. Editor: Technical Report TR-95-6, "The Global Demography Project", is now available from NCGIA, Department of Geography, University of California, Santa Barbara, CA 93106-4060, for \$25. The 75 page report + diskette describes work in which students collected over 19,000 digitized administrative polygons and associated population counts for the entire world (217 country units). It also discusses the availability of these rasters.

6. From **John Mann** to **Mac Otten**: I read your note (April, GIS Communication) about your need for a Macintosh GIS system. I recommend ARCVIEW 2 from the Environmental Systems Research Institute, Inc. (ESRI), the same folks who developed ARC/INFO. I have been recently learning the ARCVIEW 2 for Windows and found its demographic and mapping capabilities quite comparable to many of the features of the powerful ARC/INFO program. I recommend it for your use for two reasons. The price is less than \$1,000.00. The commercial price is \$995 but the government purchase price should be around \$700.00. Also, ESRI has international datasets that might include China. The nearest ESRI office to you is in Singapore at telephone number 65-735-8755 or fax 65-735-5629. The main office is in Redlands, CA at 1-800-447-9778. You can also reach them at Internet address info@esri.com.

7. ZIP codes (info from the net): In response to the query about where to find ZIP code level U.S. Census data: its out there. The files which the Bureau produces with 5-digit ZIP data are called "STF3B"; the "STF3" identifies it as a Summary Tape File 3 product (approx. 3300 variables in 200+ tables based on the long form questionnaire); the "B" is the code for the B series files, the only ones with ZIP geography (the A files have data for counties, tracts, cities, block groups, etc.-one file per state; the C file has data for the entire U.S. for larger geographic areas.) The Missouri State Census Data Center created some standard files and reports based on STF3 which are available on Internet. The only reports at the ZIP level are for Missouri, although we have the ability to generate them for anywhere in the U.S. The standard extract files (about 220 vars. derived from the 3300+) for all ZIPs in the U.S. are available from the CIESIN ftp server in Saginaw. You can ftp to [ftp.ciesin.org](ftp://ftp.ciesin.org) and then `cd /pub/census/usa cd stf cd al` (or `ca`, or `ny`, or any 2-char state code).

The ZIP files are organized by metro-area within state and are in comma-delimited format, zipped. In `/pub/census/usa/stf/0code` you'll find the SAS code that can be downloaded and used to turn the ".csv" files back into the SAS datasets from which they were originally created. (You can also get to this via www.ciesin.org--look for the U.S. Demography home page; it will take you to FTP.) If you want a quick idea of what kind of data you'll be getting use gopher to access gopher.coin.missouri.edu and select options 7,3,4,1,2 and 9 to get to Missouri ZIP tables. This will be the formatted report, what you get off the ciesin FTP server is zipped, comma-delimited data files. We also have the entire country stored on our local UNIX machine-1980 and 1990 data. We can do custom extracts, reports, etc. If you have any

interest in how ZIP code geography relates to other units such as cities, counties, metro areas, etc. you can check out the zipseq directory on the ciesin archive. These are huge files with 1 record per census block. By: John Blodgett, Urban Information Center/Office of Computing, University of Missouri - St. Louis, 8001 Natural Bridge Rd., St. Louis, Mo 63121-4499, Phone: (314) 516-6014/6000 FAX: 516-6007.

8. Census geography archives (useful for re ZIP Coding situations): "BOUNDARY FILES", BASED ON TIGER 1992 (Version 5)

- By county, boundary files for all census blocks. Each census block is identified with a unique polygon identification field (POLID).
- By county, street intersection files. For all intersections, street names and latitude-longitude of intersections are reported.
- By state, boundary files for: county, 1990 tract/block numbering areas, blockgroup, place, county subdivisions, Minor Civil Division (MCD's); and 1980 tract and MCD's (if present).
- Tract and blockgroup files are conveniently split into sub-files by Metropolitan Statistical Area (MSA/CMSA) within states.
- All boundary files in simple boundary ascii (BNA) format. Code under development for several conversion routines, including from boundary BNA format to SAS gmap datasets (SAS Institute Inc) and Arc (ESRI Inc) coverages. Code to be provided as part of the archive.
- Experimental 1% and 5% PUMA boundary files, by state, for mapping of 1990 PUMS data. PUMA boundary files are unthinned (thinned version planned for the near term).
- Complete coverage for US and all its possessions.
- BNA file format directly importable into ATLAS*GIS (SMI). Easily reconfigured for

other desktop packages.

"STANDARD EXTRACT FILES", BASED ON 1990 STF3A

- By state, files containing 225 FAFVAR (Frequently Asked For Variables) from Summary Tape File 3A (STF3A), for various geographic levels. Included are the location of the polygon centroids in latitude-longitude coordinate pairs.
- Simple, comma separated value (CSV), format.
- Data files for county, tract/bna, blockgroup, place-within-county, county subdivisions: -- all matching the corresponding TIGER-based boundary files via carefully-assigned unique polygon identification (POLID) fields.
- Tract and blockgroup files organized into MSA/CMSA (and non-metro remainder-of-state) files, matching the organization of boundary files (except in New England).
- Data files for state, metro, 5-digit ZIP code, and place-within-state (no boundary files for these geographies).
- Complete continental US coverage planned and Alaska/Hawaii. Some missing states but over 2/3 of all states are complete.
- Demographic data files split into "a" and "b" sub-files. This simplifies conversion to dBASE3 (Borland International Inc) and AGIS*GIS (SMI), as well as other desktop packages.
- Code provided to convert CSV files to SAS (SAS Institute Inc) data sets with attribute labels and formats assigned.
- Documentation online to link CSV field names to entries in "Basic Tables" reports (see Items of Interest).
- Block population counts/housing unit structures data files. By state, data files contain few variables (including population counts and housing unit structures) uniquely matched to the TIGER block boundary files via POLID.

Centroid coordinates in latitude-longitude are included.

"ENHANCED MIGRATION FILES", BASED ON STP28, COUNTY TO COUNTY

-State files with population in and population out fields on each record to facilitate net-migration calculations by age, race, sex, etc. Period of migration covers 1985 to 1990.

WHERE: ftp ftp.ciesin.org

<login in as "ftp" or "anonymous">

<send email address as password>

cd /pub/census

9. A Census contact for the Census Bureau Data Extraction System (DES). The DES is located at <http://www.census.gov/DES/welcome.html>. For data or how to questions, contact Judy Eargle, Housing and Household Economic Statistics Division, Census (301) 763-8376 or jeargle@info.census.gov.

10. Editor: I want to call your attention to the article "Effects of the choice of age-adjustment method on maps of death rates", by **Linda Pickle** and **Andrew White**, *Statistics in Medicine*, **14**, 615-627 (1995). Since many of us are involved in mapping death rates for small areas, either for stand alone or GIS purposes, the selection of method of age adjustment is critical to the correct geographic representation of the data. Rates that are indirectly adjusted, or equivalently SMRs, can be quite variable. This caution applies as well to several empirical Bayes methods which include an indirectly adjusted rate or SMR as a component. An important recommendation by the authors: map directly adjusted rates whenever possible.

11. Editor: I also want to call your attention to the article "Issues associated with the design of a national probability sample for human exposure assessment", by **Trena Ezzati-Rice** and **Robert Murphy**, in *Environmental Health Perspectives*, **103**, 55-59 (1995). They point out that although there are no current national data available on the prevalence of exposure of the U.S. population to various toxic substances, such as persons living in certain agricultural or industrial areas, national probability sample surveys and selected special studies could be designed to address these issues.

12. Editor: Global Positioning Systems (GPS) that read earth locations from space satellites, to submeter (even centimeter) accuracy, are beginning to play public health roles never before realized. In the case of one Superfund site near W. Chicago, Illinois, leftover industrial radioactive thorium is being measured using GPS and a radioactive sensor simultaneously. Transects (5 ft. apart) of the site provide continuous and precise spatial coverage for data analysis and recommended remediation.

13. Editor: I am in the process of preparing an article on GIS projects at CDC/ATSDR. If you have not responded to my prior survey of CDC/ATSDR GIS projects, please get in touch with me so that your project can be listed in the paper. I need to know 1) the disease or environmental condition you are studying, 2) the project beginning and ending dates, and 3) the geographic scale of study. Thank you for your help.

Chuck Croner, Editor, **GIS News and Information**, Office of Research and Methodology, National Center for Health Statistics, Centers for Disease Control and Prevention

Stay in GIS touch